Razonamiento aproximado

*E. Javier Salazar Resines**

INTRODUCCIÓN

Mucho se ha dicho, a favor y en contra, de la *inteligencia artificial* (IA). Parece ser que las opiniones en contra son en buena medida tangenciales y se dirigen disimuladamente en contra de las computadoras (el ataque directo suele evitarse porque el terreno de la computación se ha vuelto *casi* inexpugnable). Y el medio universitario no escapa a esta desafortunada actitud. En charlas con profesores de diversas especialidades se escuchan comentarios como éstos:

- Un cacharro de éstos, las computadoras, no puede sustituir la labor e inteligencia de un buen investigador, es algo absurdo,
- La creación no es asunto para una caja de alambres, de *chips* o de modelos matemáticos. Esto incumbe a los ingenieros o a los investigadores de las ciencias exactas.
- Los problemas que enfrentan los investigadores son de un elevado grado de complejidad y no hay

IZTAPALAPA 44
julio-diciembre de 1998
pp. 69-94

* Profesor investigador del Departamento de Economía de la Universidad Autónoma Metropolitana-Iztapalapa.

una computadora o sistema de inteligencia artificial, o de lo que sea, que pueda contender con ellos. Pretender su protagonismo en estas actividades es una utopía propia de la ciencia ficción.

- Son sistemas muy caros, fuera del alcance de los profesores y los alejan y distraen de su campo de acción.
- Sólo sirven como máquinas de escribir y en el mejor de los casos para el procesamiento estadístico o numérico. Para otros fines requieren de un entrenamiento previo muy extenso, ajeno al campo particular de trabajo.

Y otras opiniones por el estilo. Las opiniones a favor llegan a exageraciones increíbles que crean confusiones y expectativas altamente infundadas. Se sostienen señalamientos como: "el cerebro o la inteligencia de las computadoras simula ventajosamente al pensamiento humano, y eventualmente sustituirá al hombre, aunque requiere de programadores altamente capacitados e inteligentes", y muchas otras sandeces similares, buen número de ellas sutilmente apoyadas por publicidad de orden comercial.

Ambos tipos de comentarios son negativos y muy posiblemente vinculados a la ignorancia, al dogmatismo o a algún tipo de trauma psicológico.

No tiene caso polemizar sobre estas opiniones; en lugar de ello trataremos de señalar aspectos positivos que nos han conducido a orientar nuestro trabajo a nuevas opciones de razonamiento apoyándonos en las computadoras y los conceptos derivados de la inteligencia artificial, con énfasis en las denominadas estructuras del conocimiento.

Hace más de veinte años, lo que es un lapso amplio dada la evolución reciente e impresionante del área de la computación, Moles (Ferrero, 1971, 1975; Moles, 1986) apuntó que no se concibe una filosofía actual que excluya a la informática.

Ya antes se había publicado un texto sobre *Filosofía y cibernética* (Crosson, 1967). En ese libro se discutía, con bases filosóficas, cómo las computadoras, además de sus capacidades numéricas, tienen la posibilidad, a través de mecanismos inteligentes, de forzar el comportamiento racional del ser humano. Dichos mecanismos son ante todo programas apoyados por lenguajes apropiados, como *Lisp o Prolog*. Esto significa que no es la computadora la inteligente sino los programas, que son un producto humano. Las computadoras sólo son instrumentos. Las bases de *Prolog* se pueden adquirir en unos cuantos días. Si el tema es complejo o muy extenso, la tarea requerirá de más tiempo; según la complejidad pueden necesitarse meses o incluso años. Durante todo el proceso se estará en contacto con las operaciones y conceptos asociados al problema. De este modo, la complicación de la programación es la del propio problema o la de las acciones automáticas o inteligentes que se desea que realice. De manera que la complejidad es la del problema en cuestión y al programar seguramente se adquirirá un mejor entendimiento de éste. Es decir, programar es entender.

Otra referencia importante es la del doctor George (1979), quien examina temas como inteligencia artificial, método científico, explicación, incompletitud de Gödel, determinismo y otros, y anuncia otro titulo que abordará cuestiones epistemológicas, lógicas, de significado y otras más estrictamente filosóficas.

Más recientemente se han publicado libros como el de Boden (1990), que es una recopilación de destacadas contribuciones de autores como McCulloch y Pits *(cálculo de las ideas...)*, Turing *(la maquinaria de computación y la inteligencia)*, Searle *(mentes, cerebros y programas)*, Newell y Simon *(la ciencia de la computación como investigación empírica...)*, McDermott *(una crítica de la razón pura)* y algunos más. Otra publicación reciente es la de Thagard (1992). Este filósofo hace un análisis panorámico de contribuciones clave de la evolución científica (Copérnico, Kepler, Galileo, Descartes, Newton, Lavoisier, Darwin, Einstein); además, siendo filósofo desarrolla un sistema de cómputo para la inteligencia artificial (ECHO) para facilitar el análisis de los cambios conceptuales de la ciencia.

El trabajo de Thagard se basa en las aportaciones de Sowa (1984), quien discute las bases filosóficas, psicológicas y lingüísticas para crear grafos conceptuales en el razonamiento computacional y la ingeniería del conocimiento.

Muchos otros textos abordan diversos enfoques del razonamiento aproximado y generalmente cualitativo. Un libro de Fischler y Firschein (1987) presenta un panorama optimista de asuntos sobre el razonamiento, la resolución de problemas, el aprendizaje, el lenguaje, los sistemas expertos y otros temas que informan con acierto sobre asuntos centrales de la inteligencia artificial. Además de los métodos "fusos" presentados en este artículo, se han desarrollado otras dos corrientes importantes: los métodos abductivos (Neapolitan, 1990; Charniak y McDermott, 1985) a partir de inferencias causales (alternativas a las redes bayesianas, descritas en el artículo) y las redes semánticas,[1] que han tenido desarrollo, fundamentación y promoción extensos en el campo de la inteligencia artificial. En un libro relativamente reciente se hace un análisis crítico extenso sobre las propias redes semánticas, las taxonomías, la expresividad en los lenguajes naturales, la herencia y las redes, la representación del conocimiento, la clasificación, el razonamiento y los grafos, entre otros temas (Sowa, 1991). No obstante el desarrollo de las redes semánticas, el autor ha preferido su propia interpretación, en este artículo y en otros previos, sobre los "grafos" y su aplicación. Ésta es menos estricta pero también menos restrictiva y complicada en su aplicación práctica. Buena parte del material de este artículo se presentó a los alumnos del doctorado en economía en el invierno de 1995.

Finalmente es conveniente mencionar el razonamiento cualitativo que recientemente ha adquirido notoriedad en el campo de la inteligencia artificial, Kuipers (1994), analiza entre otros asuntos los conceptos del razonamiento cualitativo, de los grafos para estados discretos, las ecuaciones diferenciales, los procesos dinámicos cualitativos y otros temas de simulación y representación con información incompleta. Esto último podría no ser muy atractivo, pero el conocimiento empírico es siempre incompleto.

Con esta introducción hemos pretendido ubicar, de la mejor manera posible, los conceptos y modelos del artículo en su ámbito natural.

Aquí se presentan, de manera sucinta, conceptos fusos y probabilísticos apoyados con grafos, para sustentar el razonamiento aproximado. Como se menciona, no existen explícitamente sistemas de cómputo aplicables a este tipo de razonamiento.

El texto se acompaña de una síntesis de la lógica de clases tradicional, con el objeto de mostrar cómo la lógica fusa se construye a partir de ella.

Se abordan sucintamente dos sistemas causales: la aplicación del teorema de Bayes y las redes bayesianas.

Según se puede apreciar, la aplicación de cualquiera de los modelos descritos requiere de una cantidad extensa de información en la aplicación práctica. Se sugiere proceder por ensayo y error, interpretando los resultados en cada ciclo y ajustando los valores intrínsecos en el proceso. Esto implica hacer inferencias con información incompleta o, cuando menos, escasa.

Esta última reflexión nos lleva a la necesidad de ensayar alternativas como las anotadas en esta introducción. Con esto suponemos la poca utilidad de profundizar o de extenderse más, desde el punto de vista de la computación, en los temas abordados en el trabajo. Una opción razonable podría consistir en extender el programa *Planestr* (Salazar, 1995b) para incluir criterios más cercanos a los factores de certidumbre del *Mycin*, mencionado en el trabajo, o a las reglas de producción en procesos de búsqueda (Thorton y Du Boulay, 1992) u otros similares, más accesibles y operativos que los presentes; posiblemente la adaptación del sistema *Emicyn* (Buchanan y Shortliffe, 1984) a un programa como *Redfusa*, descrito en el trabajo, y en el caso de bases extensas, complementado con el algoritmo *Rete* (González y Dankel, 1993) para decidir el acoplamiento óptimo de todas las reglas y hechos, mediante redes de patrones y de valores asignados a las variables.

Notas sobre redes e incertidumbre para el razonamiento aproximado

Con base en los factores de certidumbre del *Mycin* (Shortliffe, 1976), que a su vez partieron de las fórmulas de Bayes, Marcellus (1989), desarrolló un procedimiento accesible para el razonamiento inferencial (en red), con la propagación de valores. El método se usa en los programas *Grafuz* y *Redecis* (Salazar, 1994 y 1994a) y se introduce en el artículo sobre *incertidumbre* (Salazar, 1995).

Marcellus parte de cuatro cuestionamientos básicos:

1. ¿Cómo cuantificar nuestra creencia de que una "pieza" de evidencia es o no verdadera?
2. ¿Qué tanto apoya una evidencia a una conclusión?
3. Si hay evidencias alternativas para una misma conclusión, ¿cómo debemos operarlas?
4. ¿Cómo operar el razonamiento incierto y concatenado?

En el apéndice del artículo sobre la incertidumbre se menciona cómo Shortliffe propuso sus *factores de certidumbre* (FC) a partir de la fórmula de Bayes, y de allí Marcellus pasa a las redes como método de inferencia concatenada y expone, con alguna extensión, las limitaciones y complicaciones implícitas en la aplicación práctica.

La expresión de probabilidad condicional

$$p(a/b) = \frac{p(a\&b)}{p(b)} \quad (\& \text{ en lugar de } \wedge)$$

es el origen de la fórmula de Bayes, que a su vez suele expresarse en sus dos alternativas básicas (Salazar, 1990):

a) La probabilidad de H en la presencia de la evidencia E:

$$p(H/E) = \frac{p(E/H)p(H)}{p(E/H)p(H) + p(E/\neg H)p(\neg H)}$$

b) La probabilidad de H en la ausencia de E ($\neg E$):

$$p(H/\neg E) = \frac{p(\neg E/H)p(H)}{p(\neg E/H)p(H) + p(\neg E/\neg H)p(\neg H)}$$

La aplicación directa de la fórmula de Bayes requiere de una cantidad extensa de información previa. En algunas áreas del conocimiento (o de "sistemas expertos")[2] se tienen bases muy amplias de información (como en el caso del diagnóstico médico). Se pueden inventar los datos pero, cuando esto sucede, el rigor matemático deja de tener sentido.

La situación se complica si hay varias evidencias que concurren a una sola conclusión. En el caso de evidencias conjuntivas cabe hacer la composición sugerida en el artículo "Incertidumbre", lo que implica que los eventos de la composición sean independientes entre sí. En las fórmulas anteriores habrá de sustituirse E por la conjunción de otros eventos: E1, E2, ..., En.

Es decir:

E por E1&E2& ... &En

en todas las apariciones de E. Puede fácilmente imaginarse la complicación de las expresiones y la amplia imposibilidad de obtener la información de las probabilidades de las conjunciones y de las negaciones de tales conjunciones.

El proceso es aún más complejo en la sustitución de disyunciones:

E por E1 v E2 v ... v En

donde, además, habrá que contabilizar las conjunciones ya incluidas y sus negaciones.

Se tendrían que obtener las probabilidades de

p(H/E1&E2&...&En)

para la evidencia conjuntiva y

p(H/E1vE2v...vEn)

para la evidencia disyuntiva.

En casos extremos de cálculo —no necesariamente del razonamiento—

se tendrían combinaciones de disyunciones, conjunciones y negaciones, por ejemplo:

p(H/(E1&E2)&(¬E3vE4))

que con las fórmulas para el teorema de Bayes se llega a complicaciones excesivas en su aplicación práctica.

El rigor matemático y la aplicabilidad

Para la aplicación de redes en el razonamiento incierto se definen los factores de certidumbre, mencionados en el artículo sobre la incertidumbre.

La probabilidad simultánea (o conjuntiva) de dos eventos a y b es, a partir de la probabilidad condicional:

p(a & b) = p(a/b)p(b) = p(b/a)p(a)

donde se puede expresar que

si a ∈ A y b ∈ B

o sea que a y b están en el producto cartesiano

AXB

En términos de las definiciones de Zadeh,[3] en cuanto a las posibilidades conjuntas de a y b, a está en A *si* b está en B.

En esta expresión, suponiendo que

- el *factor de certidumbre de la evidencia* es FC(e) ≈ p(e), y que
- el *factor de certidumbre de la inferencia* es

  FC(i) ≈ p(c/e)

es decir que la certeza de que si la evidencia es verdadera, y la certeza de la inferencia (implicación) es FC(i), entonces la certeza de la conclusión estará determinada por

FC(c) = FC(e) x FC(i), o sea

FC (conclusión) = FC (evidencia) x FC (implicación)

que son las expresiones desarrolladas por Shortliffe en *Mycin* (o *Emycin*).

Usando "ct" para denotar la certidumbre:

ct (conclusión) = ct (evidencia) x ct (implicación)

Zadeh y Shortliffe usan diversos valores para la *ct de la evidencia*, según sean sus combinaciones:

- Para la evidencia *conjuntiva* (&)

  ct(e1 & e2) = min[ct(e1), ct(e2)]
- Para la evidencia *disyuntiva* (v)

  ct(e1 v e2) = máx[ct(e1), ct(e2)]

Es usual expresar las inferencias con evidencias disyuntivas, mediante inferencias separadas que concurren a una misma conclusión:

r1: Si (e1) entonces (c)   ct (conclusión) = .8

r2: Si (e2) entonces (c)   ct (conclusión) = .7

En esta situación el ct (total) será:

ct1 (total) = ct(r1) + ct(r2) - ct(r1) x ct(r2)

Para las reglas anteriores:

ct1 (total) = 0.8 + 0.7 - 0.8 x 0.7 = 1.5 - 0.56 = 0.94

Si hubiera otra regla:

r3: Si (e3) entonces (c)   ct (conclusión) = .6

se tendría:

ct2(total) = ct(r1) + ct(r2) + ct(r3) - ct(r1) x ct(r2) - ct(r1) x ct(r3) - ct(r2) x ct(r3) + ct(r1) x ct(r2) x ct(r3)

numéricamente:

ct2(total) = .8 + .7 + .6 - .8 x .7 - .8 x .6 - .7 x .6 + 8 x .7 x .6
= 2.1 - 0.56 - 0.48 - 0.42 = 2.1 - 1.46 + .336
= 0.976

Es fácil probar que

ct2 (total) = ct1 (total) + ct(r3) - ct(r3) x ct1(total)

numéricamente:

ct2(total) = 0.94 + 0.6 + 0.6 x 0.94 = 1.54 - 0.564 = 0.976

Esta nueva expresión simplifica los cálculos (sobre todo en la programación).

Extensión de la certidumbre[4]

De manera similar a los desarrollos de Naylor y de *Expincer*, ya mencionados, se pueden extender las certidumbres "ct" para incluir evidencias negativas:[5]

ct(not e) = -ct(e)

Así se tienen los valores:

-1 cuando la evidencia o conclusión es *falsa*

+1 cuando la evidencia o conclusión es *verdadera (o cierta)*,

0 cuando no se tenga conocimiento del evento.

o valores intermedios entre -1 a +1.

Cuando hay más de una ct en dos reglas alternas se toman las siguientes opciones:

- si ct1 y ct2 son *ambas positivas*

  ct(total) = ct1 + ct2 - ct1 x ct2
- si ct1 y ct2 son *ambas negativas*

  ct(total) = ct1 + ct2 + ct1 x ct2
- si *una es positiva* y la *otra negativa*

  $$\text{ct(total)} = \frac{\text{ct1} + \text{ct2}}{1 - \min(\text{abs(ct1)}, \text{abs(ct2)})}$$
- si *una es +1 y la otra -1*

  ct(total) = 0

De los esquemas básicos para la inferencia en el artículo "Incertidumbre" se pueden construir las redes para los programas *Grafuz y Redecis*.

Para ambos programas se incluyen ejemplos de bases de inferencias y, además, las bases con símbolos mínimos para *ejercitar* aplicaciones.

Los programas incluyen las combinaciones de inferencias AND, OR y NOT. Las bases de *Grafuz* son distintas, más elementales, que las de *Redecis*.

Introducción a los conjuntos fusos

En la literatura sobre los conjuntos "fuzzy" suele traducirse a éstos como "conjuntos borrosos". Existe en español el término "fusco" (Menéndez, 1965), equivalente al "fuzzy" del inglés, y que incluso se usa como desinencia en algunas palabras: confuso o difuso; este término hace alusión a algo que es *oscuro*. Por ello, decidimos emplear el término *fuso* como traducción de "fuzzy".

Se utiliza otro término para referirse a los conjuntos de la lógica estándar: "conjuntos crisp"; esto también podría tener un equivalente en español: "conjuntos crespos o crispados". Se emplea esta denominación para dar cuenta de una representación gráfica de los valores 1 (verdadero) y 0 (falso) que toman los elementos de un conjunto:[6]

- 1, si el elemento pertenece o es miembro de una clase, por ejemplo, si un elemento x es miembro de una clase A, $x \in A$, y

- 0, si el elemento no es miembro de una clase, $x \notin A$ (x no pertenece a la clase A).

En todo caso, lo importante es que en un universo dado los elementos pueden asumir los valores 0 o 1. De manera funcional:

$$f_A(x) = \begin{cases} 1 \text{ si y sólo si } x \in A \\ 0 \text{ si y sólo si } x \notin A \end{cases}$$

En el apéndice A se incluyen propiedades y operaciones booleanas de la lógica de clases (o conjuntos) estándar.

Funciones booleanas disyuntivas

En un artículo del autor (Salazar, 1993) se introdujeron las llamadas *funciones booleanas normales disyuntivas*, para la determinación de las trayectorias en grafos y redes (implicantes y diagramas de Venn); en ingeniería se emplean en el diseño de circuitos lógicos para computadoras (Nagle, Carrol e Irwin, 1975). Para estas tareas se suelen representar las operaciones básicas mediante las denominadas *compuertas lógicas*, como las mostradas en la siguiente figura.

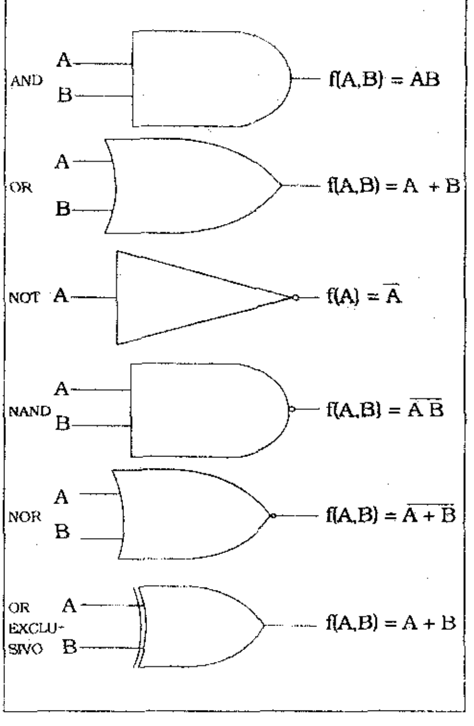
El diseño lógico básico implica sólo combinaciones de estas *compuertas*. Un circuito consiste en combinaciones, como se muestra en la figura de la siguiente página:

La función f(A,B,C) a la salida del circuito toma los valores 0 y 1, según sean los valores de A, B y C de la entrada y las combinaciones de compuertas.

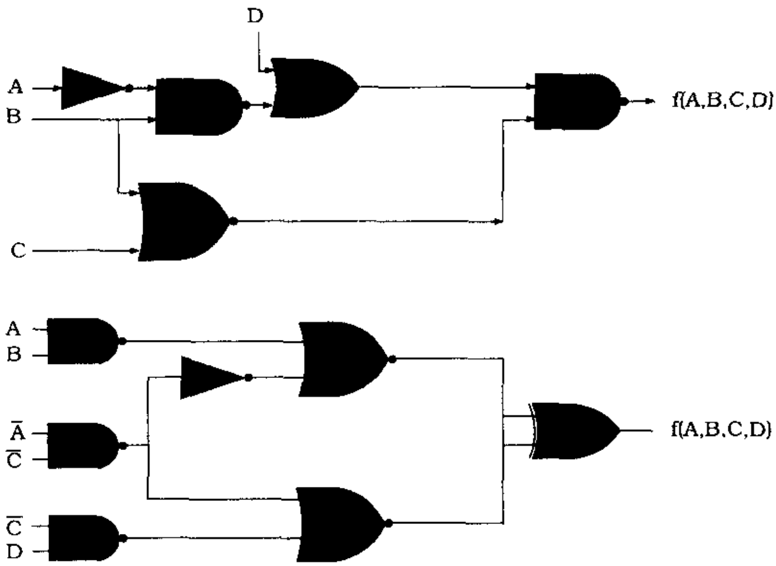
Su aplicabilidad se puede extender a las ciencias sociales, entre otras razones porque la lógica no es privativa de alguna área técnica particular. Un entusiasta grupo de profesores de la UAM (Iztapalapa y Xochimilco) produjo notables trabajos de investigación en el campo de las ciencias sociales (Salazar, Peñalva y Leal, 1993).

En un texto de lógica (Salazar, 1990) se presentaron casos de funciones disyuntivas con combinaciones de valores de probabilidad. Este último trabajo, como muchos otros que se ofrecen como

Compuertas lógicas



*Circuitos con compuertas lógicas:**



* Nagle, *An introduction to Computer Logic.*

herramienta a los investigadores para el razonamiento aproximado, son ensayos inconclusos que distan de tener una aplicabilidad práctica. Sólo abren caminos para futuras aplicaciones. Abundan este tipo de intentos en la literatura sobre los expertos y la inteligencia artificiales.

Otro ensayo que vale la pena destacar, que opera con expresiones booleanas elementales y que se aplica de manera confusa al cálculo de probabilidades, es la de Ursic Computing (1993). En este sistema se operan expresiones conjuntivas, disyuntivas y bayesianas. Opera con valores numéricos sobre un diagrama de Venn; por concatenación puede incluir funciones de cierta complejidad. Aunque denomina a su trabajo como *razonamiento fuso*, el proceso es más bien probabilístico.

A estos ensayos habrá que agregar los "expertos artificiales" de índole comercial, cuya única meta es vender el producto. Los usuarios deben estar atentos para no caer en las promociones, por demás exageradas y extensas, de los creadores de tales productos (por ejemplo, la programación orientada a objetos, "redes neurales"; "expertos artificiales" comerciales y otros temas que suele confundir a los investigadores y usuarios potenciales).

Nociones sobre conjuntos fusos

Supuestamente los conjuntos fusos contienden con algunos tipos y grados de incertidumbre. Ésta se refiere esencialmente a la que surge en expresiones del lenguaje cotidiano (que en última instancia es el empleado por los investigadores como metalenguaje). Esto, independientemente de la terminología de cada campo particular. Pero hay cuestiones derivadas de la incertidumbre que no se originan en el lenguaje, por ejemplo la asociada a la complejidad, la aleatoreidad de los datos, la emotividad o los hechos no bien definidos (Salazar, 1995).

Aunque nos restrinjamos a la incertidumbre originada en el lenguaje natural no estamos con ello abordando un tema intrascendente; sólo basta considerar que por falta de claridad en sus expresiones, vaguedades o imprecisiones (Cox, 1994), los científicos, no se escapan a las trampas y vericuetos del lenguaje natural. La fusividad (oscuridad) está vinculada más estrechamente con la imprecisión y la vaguedad. Otras acepciones son muy cuestionables y de alcance difícil de definir.

Algunos ejemplos de expresiones vagas y usuales son:

"Medio lleno o medio vacío"
"La situación es terrible"
"La temperatura es más bien alta"
"La inflación crece rápidamente"
"Las investigaciones extensas se desarrollan con lentitud"
"Si Pedro es alto entonces es pesado"

Y muchas de las expresiones del lenguaje cotidiano.

A diferencia de la lógica tradicional, en donde la verdad de una aseveración se hace corresponder con la pertenencia (o membresía) de un elemento a una clase: verdadero si pertenece y falso si no pertenece (un animal x es o no es un mamífero), en la lógica fusa la pertenencia (o membresía) tiene un nivel o grado. Éste se define mediante una curva de membresía. Así, un personaje x puede ser "alto" en un cierto grado; su "altura" va a depender de lo que se entienda o defina como tal. En la lógica tradicional se entiende que la definición es contextual o teleológica (depende del asunto o de la finalidad). Dentro de la lógica fusa, debido a la innumerable cantidad de interpretaciones que se le han dado a las ideas originales de Zadeh (creador de los conjuntos fusos), no aparece muy clara la definición de un universo del discurso para cada aseveración.

Donde mejor se definen estos conceptos es en los trabajos de Klir (Klir y Folger, 1988) y de Cox (1994); algunos de estos conceptos se proporcionan en la siguiente ilustración.

Resumen de operaciones básicas en conjuntos fusos

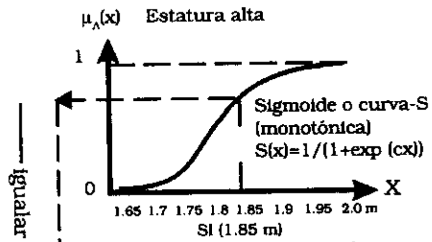


figura 1

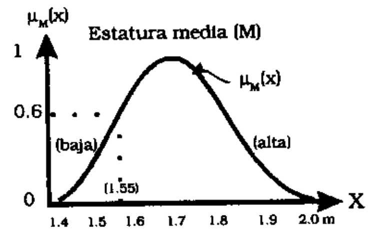


figura 2

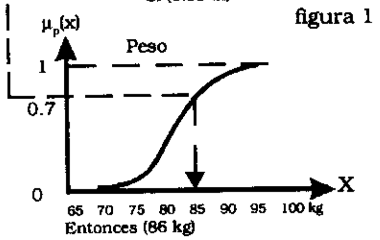


figura 3

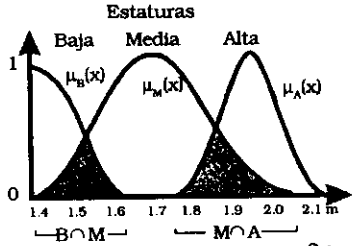


figura 4

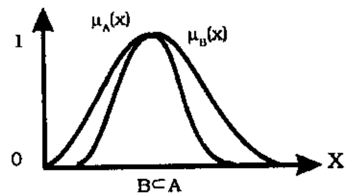


figura 5

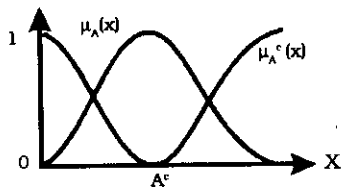
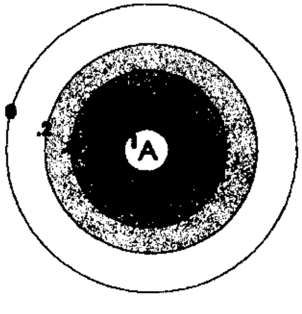


figura 6



Venn Fuso

figura 7

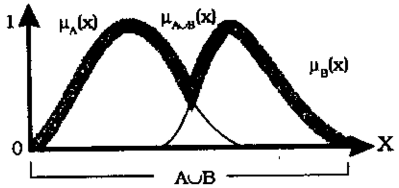


figura 8

Con la letra griega μ se indica la medida de la "membresía". Por ejemplo para la *estatura alta* se usa una curva S (sigmoide),[7] asintótica a cero para valores muy bajos de m y asintótica a 1 para valores superiores a los 2 m. Entre las figuras 1 y 3 se ilustra un caso simple de inferencia (dentro de algún contexto): Si un sujeto tiene una estatura de 1.85 m de estatura con una membresía de 0.7, entonces tendrá un peso de 85 kg (para la misma membresía).

En la figura 2 se ilustra el concepto vago de *estatura media*. No puede discernirse sobre este concepto sin más. Habrá que definir el universo (que incluya el propósito, la población, el lugar y quizá la fecha). Este concepto, como muchos otros similares mencionados por Cox, se asocian a situaciones con posibilidades (membresías) con variación *no-monotónica;* no en un caso, como el anterior, en donde la membresía crece continuamente al aumentar la estatura. Si el máximo de la estatura media es, por decir algo, 1.68 m, abajo de este valor se puede tener una estatura no *muy media;* en cierto límite puede ser claramente *baja,* y arriba también puede ser *no muy media*, más bien alta. Algunos conceptos vinculados con este tipo de variación son, por ejemplo: "algunos", "unos cuantos", "alrededor de", "en la vecindad de". Este tipo de variación puede tener expresiones matemáticas[8] de mayor o menor complejidad para expresar la membresía (o grado de posibilidad).

En la figura 4 se muestra la superposición de varias curvas: baja, media y alta para el concepto vago *estatura*. Las zonas de intersección indican la indefinición de conceptos. En la figura 5 se esquematiza la curva para la membresía de un conjunto A y del subconjunto B de A. En la figura 6 se muestran las curvas de membresía para un conjunto A y su complemento A^c.

En la figura 7 se esquematiza un diagrama de Venn fuso. Los círculos concéntricos indican curvas de nivel asociadas a diversos grados de membresía.

En la figura 8 se muestra la curva de membresía asociada al conjunto unión de los conjuntos A y B.

En ocasiones, puede disponerse de los valores numéricos de la membresía para distintas opciones marcadas en el eje x. Se suelen indicar con la expresión disyuntiva:

$$1.4/0 + 1.5/0.2 + 1.6/0.7 + 1.7/1 + 1.8/0.7 + 1.9/0.2 + 2.0/0$$

según el caso, se puede aproximar el valor de membresía.

Con los criterios de Zadeh (Terano, Asai y Sugeno, 1992):

$a \wedge b$ para el *mínimo* de a y b,

$a \vee b$ para el *máximo* de a y b,

se establecen

$\mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x)$, para la unión de A y B,

$\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x)$, para la intersección de A y B,

$\mu_{\bar{A}}(x) = 1 - \mu_A(x)$, para el complemento de A.

Como la lógica fusa sólo es una extensión de la lógica tradicional, se definen otras operaciones similares a las de las clases en los conjuntos fusos:

idempotencia, leyes de De Morgan, etcétera. Algunas con alteraciones intuitivamente inteligibles de la forma:

- equivalencia: A = B si y sólo si $\mu_A(x) = \mu_B(x)$,
- inclusión: $A \subset B$ si y sólo si $\mu_A(x) \leq \mu_B(x)$
- $A \cup \bar{A} \neq 1$ (1 universo)
- $A \cap \bar{A} \neq \phi$ (ϕ vacío)
- y otras similares.

De acuerdo con los criterios de Zadeh, se puede formar una tabla (abajo) para la combinación de dos variables x, y, separadas por 'and' y 'or'.

Con estos criterios y otros similares podremos plantear un ensayo de *razonamiento aproximado* que no se expone claramente en la literatura existente sobre conjuntos y lógica fusos.

Bases para el razonamiento aproximado

En la figura de la siguiente página se bosquejan algunos conceptos para la construcción de bases inferenciales para el razonamiento aproximado.

En la parte superior se tienen tres figuras asociadas a las curvas de membresía para la temperatura de un motor caliente. En la primera curva, el eje vertical muestra dos casos extremos: *verdadero* y *falso*. En el intervalo medio se tienen valores como: *altamente verdadero, escasamente verdadero* y *usualmente falso*.

En la parte media se tiene una aproximación lineal de la primera curva. La tercera curva es la representación numérica del eje de membresía, de 0 a 1.

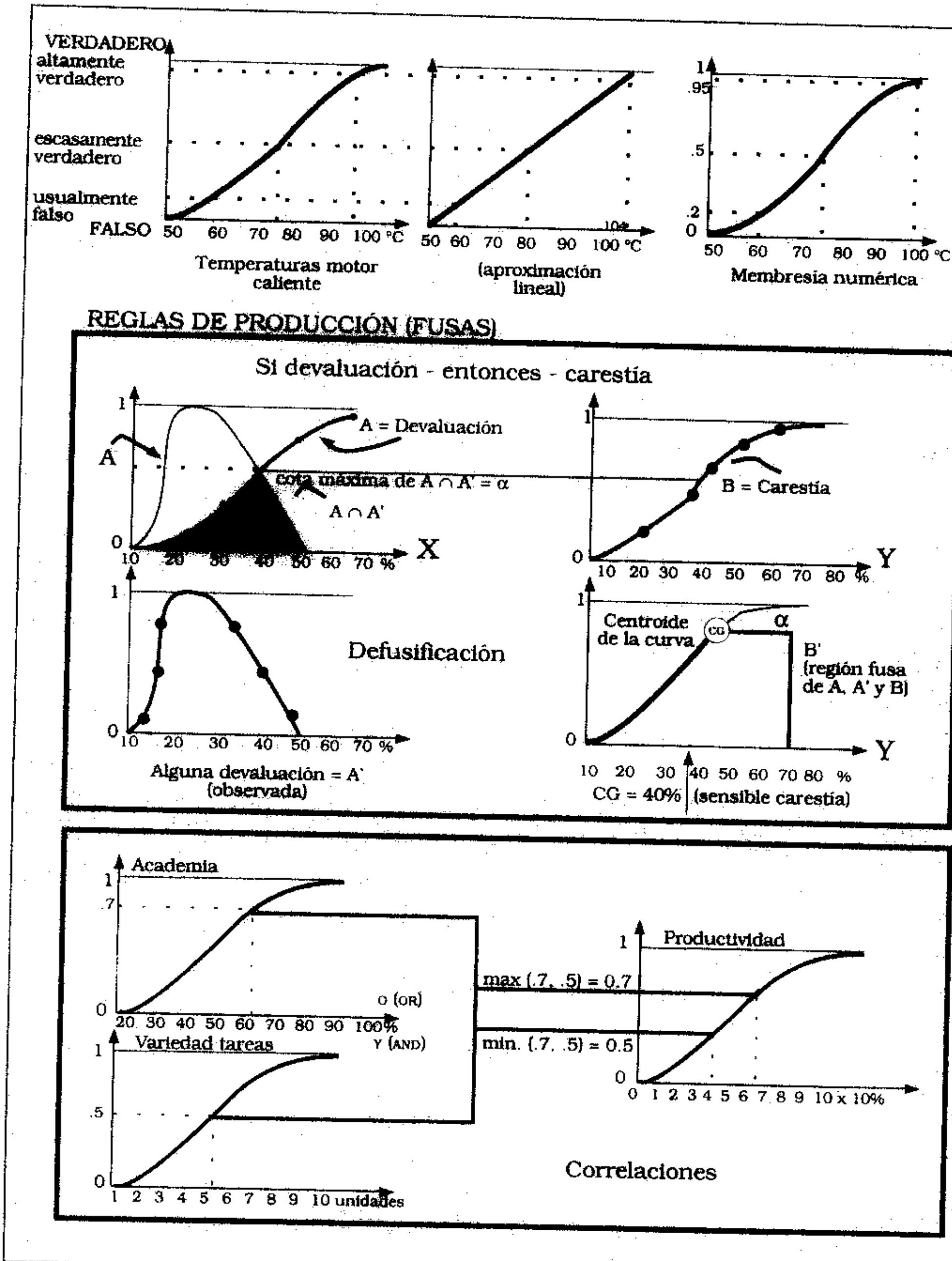
En los siguientes cuadros de la figura se ilustran las reglas de producción de tipo *fuso*.

En *defusificación* se tiene una regla del tipo *si-entonces*; corresponde a una inferencia como: "si hay devaluación, entonces habrá carestía".

El modus ponens de la lógica crespa requeriría de la afirmación: "hay devaluación" y se concluiría "habrá carestía".

En una inferencia fusa se podría afirmar: "hay alguna devaluación", según observaciones dadas en la curva abajo. Ninguna de las curvas es, necesariamente, una función analítica (matemática).

	AND					OR				
x: :y	0.00	0.25	0.50	0.75	1.00	0.00	0.25	0.50	0.75	1.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.75	1.00
0.25	0.00	0.25	0.25	0.25	0.25	0.25	0.25	0.50	0.75	1.00
0.50	0.00	0.25	0.50	0.50	0.50	0.50	0.50	0.50	0.75	1.00
0.75	0.00	0.25	0.50	0.75	0.75	0.75	0.75	0.75	0.75	1.00
1.00	0.00	0.25	0.50	0.75	1.00	1.00	1.00	1.00	1.00	1.00


VERDADERO
altamente verdadero
escasamente verdadero
usualmente falso
FALSO
50 60 70 80 90 100 °C
Temperaturas motor caliente
50 60 70 80 90 100 °C
(aproximación lineal)
1
.95
.5
.2
0
50 60 70 80 90 100 °C
Membresia numérica
REGLAS DE PRODUCCIÓN (FUSAS)
Si devaluación - entonces - carestía
A = Devaluación
A'
cota máxima de A ∩ A' = α
A ∩ A'
10 20 30 40 50 60 70 %
X
B = Carestía
10 20 30 40 50 60 70 80 %
Y
Defusificación
Alguna devaluación = A' (observada)
Centroide de la curva
CG
α
B' (región fusa de A, A' y B)
10 20 30 40 50 60 70 80 %
CG = 40%
(sensible carestía)
Academia
.7
20 30 40 50 60 70 80 90 100%
O (OR)
Y (AND)
Variedad tareas
.5
1 2 3 4 5 6 7 8 9 10 unidades
Productividad
max (.7, .5) = 0.7
min. (.7, .5) = 0.5
0 1 2 3 4 5 6 7 8 9 10 x 10%
Correlaciones

En este caso debe conocerse la curva de membresía de la conclusión (B). Superponiendo la curva de observaciones con la de devaluaciones (A) se tiene una intersección $A \cap A'$. Con el valor máximo dado por la cota a y llevado al conjunto B se determina un conjunto fuso B'. Del *centroide* de la curva se obtiene un valor *pesado* para el resultado (conclusión). La posición de CG se puede calcular con:

$D_{CG} = \Sigma y_i \mu_i(y_i)/\Sigma \mu_i(y_i)$ (i=1 a n, y_i en el dominio de B' : Y)

La inferencia fusa tiene la forma:

> "*Si hay devaluación* entonces *habrá carestía.*
> *Hay alguna devaluación*
> Entonces
> *habrá una sensible carestía*"

Este tipo de inferencias, características en el lenguaje *cotidiano*, no son posibles en la lógica tradicional.

En particular, el conjunto de inferencias fusas puede expresarse como el conjunto de relaciones R ($A \rightarrow B$) sobre el dominio dado por el producto cartesiano:

$X \times Y = \{ (x,y) : x \in X \,\&\, y \in Y \}$

El programa *Redecís*, descrito en el artículo *Incertidumbre*, se presta a una primera aproximación para redes fusas. Se pueden introducir intervalos de $\mu(x)$, cada uno de los cuales corresponda a un nivel de conclusión. Las reglas podrán asociarse a las reglas de correlación del cuadro inferior de la figura anterior. Las operaciones son similares a las efectuadas con los factores de certidumbre. Habrían de calcularse los valores de $\mu(x)$ para cada nodo inicial y terminal de la red.

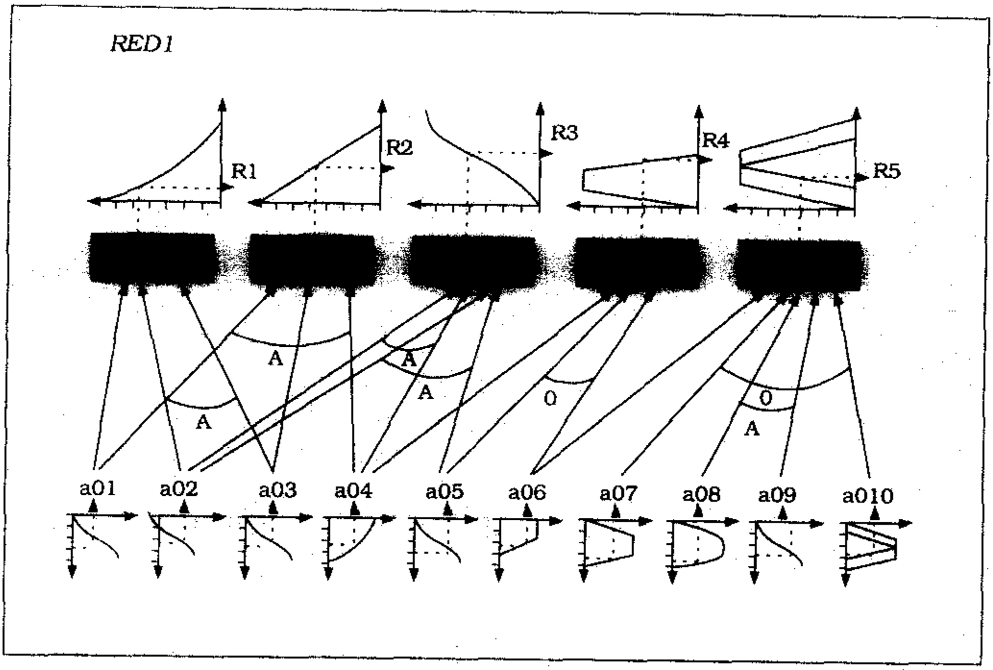
Para la adaptación se elaboró el programa *Redfusa* (Salazar, 1995c), con el cual es posible elaborar redes fusas como la que se esquematiza en la siguiente figura. La red *REDI* consta de un conjunto de nodos iniciales: *a01*, *a02*, ..., *a10* y un conjunto de nodos finales: *k01*, *k02*, *k03*, *k04* y *k05*.

Estos últimos pueden corresponder a conceptos que no interese especificar. Se creó la base de datos correspondiente, que se detalla más adelante.

La base *REDI* (siguiente página) tiene la siguiente forma:

Las reglas (de inferencia) contienen 8 campos. El primero indica el tipo de inferencia: s (simple), a ('and': conjunción) y o ('or': disyunción). El segundo campo indica la *conclusión* de cada inferencia según las premisas de los campos cuarto y quinto; éstas están precedidas por 'pos' o 'neg', según que operen con signos positivo o negativo. En las inferencias simples (s) en los campos quinto y sexto aparecerá la palabra 'muda', que indica que no hay términos ni signos en ellas. Los factores *iniciales* a01, ..., a10 se emplean según se puede examinar en las reglas, y con sus valores *iniciales* se determinan las conclusiones *finales* k01, k02 y k03. Los valores de los factores iniciales y finales se obtienen de curvas de membresía.

Los valores a0i y k0r pueden fácilmente sustituirse por frases específicas para una red de aplicaciones, mediante un procesador de textos.



Las conclusiones (R1, ..., R5 de la figura) se asocian a algún asunto específico según lo determine el grado de membresía.

El último campo de las reglas *imp* es 1. Los valores distintos a éste se emplean en el caso de factores de certidumbre (como en *Redecis*).

Para el trazo inicial de la red se puede emplear el programa *Red* y posteriormente elaborar una base como la anterior. Con un graficador puede afinarse el trazo de la red original, como se hizo con RED1.

Las curvas de membresía de los valores iniciales (a01, ..., a010) y de las conclusiones (k01, ..., k05) corresponden a un conocimiento específico.

Relaciones, grafos y matrices fusos

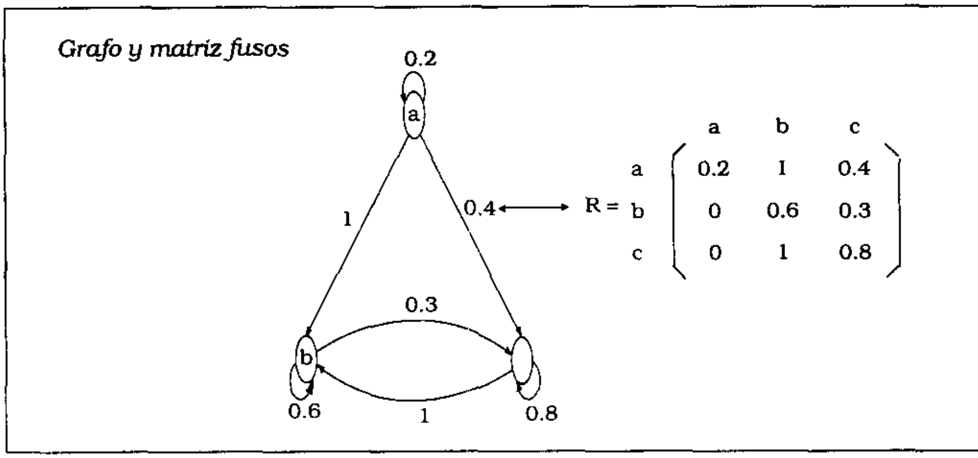
$$\text{Si } X = \{x_1, x_2, \ldots, x_m\} \text{ y}$$
$$Y = \{y_1, y_2, \ldots, y_n\}$$

se puede definir una relación fusa en el producto cartesiano X x Y, que puede expresarse con una matriz de mxn, en donde cada valor expresa un grado de membresía

$$\{ x_i, y_j \}$$

se tiene así una matriz para R

$$R = \begin{vmatrix} \mu_R(x1,y1) & \mu_R(x1,y2) & \ldots & \mu_R(x1,yn) \\ \mu_R(x2,y1) & \mu_R(x2,y2) & \ldots & \mu_R(x2,yn) \\ \cdots & \cdots & \cdots & \cdots \\ \mu_R(xm,y1) & \mu_R(xm,y2) & \ldots & \mu_R(xm,yn) \end{vmatrix}$$

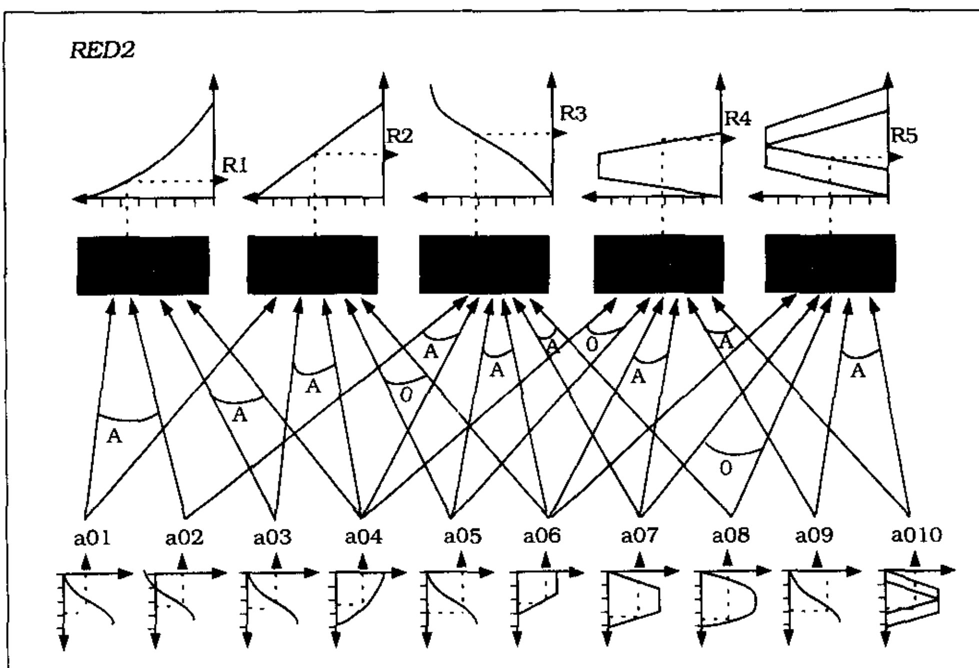


Grafo y matriz fusos

En la figura se muestra la relación

R = 0.2/(a,a) + 1/(a,b) + 0.4/(a,c) + 0.6/(b,b) + 0.3/(b,c) + 1/(c,b) + 0.8/(c,c),

todos los valores de $R{:}X \times Y \to [0,1]$. Las relaciones y los grafos son conjuntos de parejas ordenadas sobre los que se pueden definir diversas operaciones entre clases[9] (apéndice B).



RED2

NOTAS ADICIONALES

En la figura RED2 se ilustra un caso más extenso de red fusa, introducida desde el propio programa *Redfusa*.

Aunque este caso ya es más amplio, es aún muy reducido en cuanto a factores iniciales: a01, ..., *a010 y* conclusiones (membresías) *k01 .... k05*. Como antes los factores iniciales, deben provenir de asuntos concretos y culminan en membresías que se deben asociar a resultados *R1, . . ., R5*.

Por ejemplo, si se tratara de diagnósticos médicos, a01 podría corresponder a *algo de temperatura (0.45)*, a02 a un *cierto malestar de garganta (0.6)*, a03 a *flujo nasal regular (0.6)*, a04 a *malestar general más que regular (0.7)*,... y así para los demás valores iniciales. Con estos valores, operando con el programa se tendría *k01 = 0. 78*. Y un médico podría asociar a *k01* el resultado *R1 = 'quizá un poco de gripe'*. No afirmamos que las curvas de membresía para k01, ..., k05, k01, ..., k05 y R1, ..., R5 tengan una forma y significado correctos para un diagnóstico médico real, las señalamos sólo como posibles aplicaciones. Es importante advertir que no basta con *imaginar o intuir* valores para a01, ..., a10 o para k01, ..., k05, habrá que conocer (llevar registros) de las variaciones de los factores iniciales (síntomas) y de los resultados *R1* para k01, R2 para k02, ..., R5 para k05. En pocas palabras, se debe tener un conocimiento real y detallado del asunto que se maneja. Parece que así debe ser, ¿o no?

En todo caso, podremos suponer la variación de algunos factores, recurriendo a la experiencia y al sentido común. Por ejemplo, si la curva asociada a k01 corresponde a diversos grados de gripe, no puede ser que tenga la forma de la figura, en donde entre mayor sea k01 (de 0 a 1) menor es la posibilidad de esta enfermedad. Diríamos que opera al revés.

En cambio, si se refiriera a los factores positivos a01, a02, a03, a04, que inciden en la producción, por ejemplo mayor desempeño en el trabajo, mejores condiciones laborales, más alta retribución económica, reconocimiento más amplio y oportuno, y k01 estuviera asociado a la permanencia del trabajo, entre más factores positivos haya mayor es su posibilidad. Esto parece aceptable. Aún así no puede, a primera vista, decirse que la variación sea la ilustrada (considerarse una primera aproximación lineal).

Puede alegarse que no siempre se tienen datos exhaustivos para hacer, estas correspondencias. Lo aceptable será una combinación del sentido común y la experiencia con datos más precisos. Es decir, suponer *razonablemente (?)* algunas curvas y combinarlas con otras que se basen en datos recabados de la realidad (mediciones, encuestas, datos estadísticos confiables, etcétera).

Estas aproximaciones conllevan operaciones de ensayo y error. Se procesa con valores supuestos o reales y se juzgan los resultados. Las curvas de resultados también serán sujetas a escrutinio. Estas últimas son determinantes en los procesos. Aun si se tienen factores iniciales bien determinados, las asociaciones finales pueden ser un fiasco. Por ello, en los trabajos de campo no pueden diseñarse encuestas u otras formas de recopilación de datos que excluyan las metas u objetivos. Las etapas finales son *interpretaciones* sobre todas las combinaciones del proceso. Incluso puede ocurrir que con datos burdos pueda haber asociaciones finales aceptables; pero no en sentido opuesto. Estas interpretaciones finales suelen desempeñar un papel crucial, por ejemplo en el caso del riesgo sísmico, en las predicciones de crisis o en muchos otros casos en donde se *opera con información incompleta* o de baja confiabilidad, pero en los cuales las implicaciones son, o pueden ser, de relevancia considerable.

Fusividad y probabilidad

Se ha presentado una polémica entre probabilistas y fusistas, sobre todo en los procesos de razonamiento con incertidumbre. El fusista Kosko (1992, 1993) señala cómo los probabilistas bayesianos E. T. Jaynes y D. Lindley opinan que sólo la probabilidad es adecuada para el procesamiento de problemas con *Incertidumbre* y todos los demás son inadecuados. Para rebatir semejante afirmación Kosko ideó un ejemplo geométrico. Se observa la figura de algo que parece ser un círculo y se hace la pregunta: ¿"es esto un círculo"? Para obviar toda la discusión de Kosko supongamos que alguien, suspicazmente, responde: "probablemente sí". No se responde en términos de la probabilidad matemática. No se puede establecer la probabilidad de la circularidad de la figura. Ésta es una propiedad fusa, ambigua. En cambio, dentro de los calificativos fusos sí hay la posibilidad de afirmar que tal figura es algo circular, o más o menos circular, o se trata de un círculo fuso en mayor o menor grado. Sería ilógico decir que se tiene, según sus irregularidades, un círculo con una determinada probabilidad.

Esta es la discusión clásica e inútil de los teóricos puros. A quienes nos interesa sólo resolver problemas de incertidumbre no nos concierne ni nos interesa concederle la razón a alguno de ambos bandos. Después de todo, la lógica fusa y la probabilidad bayesiana tienen ambas serias limitantes. Y no contamos con herramientas mejores. El razonamiento fuso no está completamente establecido, hay demasiadas reglas y criterios enfocados sólo a problemas particulares, según puede apreciarse en las referencias antes anotadas. Por otro lado, el razonamiento bayesiano estricto, como ya se mencionó, también es muy inoperante en su aplicación práctica.

Redes bayesianas

Con el programa *Redfusa*, antes examinado, se vieron redes que excluyen los nodos intermedios. Bajo algún criterio podrían incluirse. Por ejemplo, suponiendo para los nodos subsecuentes a un nodo intermedio el valor que se obtiene de las ramas que concurran a dicho nodo, después de aceptar la interpretación asociada a su grado de membresía. En este caso, después de aceptar la interpretación opera como los nodos iniciales, en RED1 y RED2. En la siguiente figura se ilustra esta propuesta.

No existe información clara al respecto, como en factores de certidumbre.

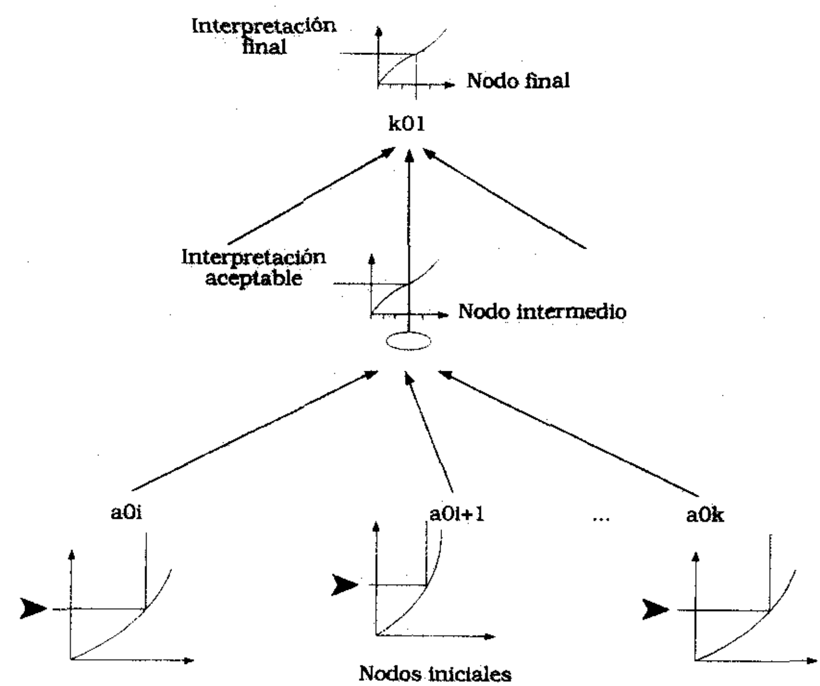
De las ramas que concurren al nodo intermedio (de a0i, a0i+1, ... a0k), que pueden ser simples, 'AND' u 'OR', se determina el valor de vi0. Con este valor se verifica alguna interpretación, de acuerdo con una curva de membresía.

Si la interpretación es aceptable, se prolonga la rama que emana del nodo intermedio con el valor vi0.

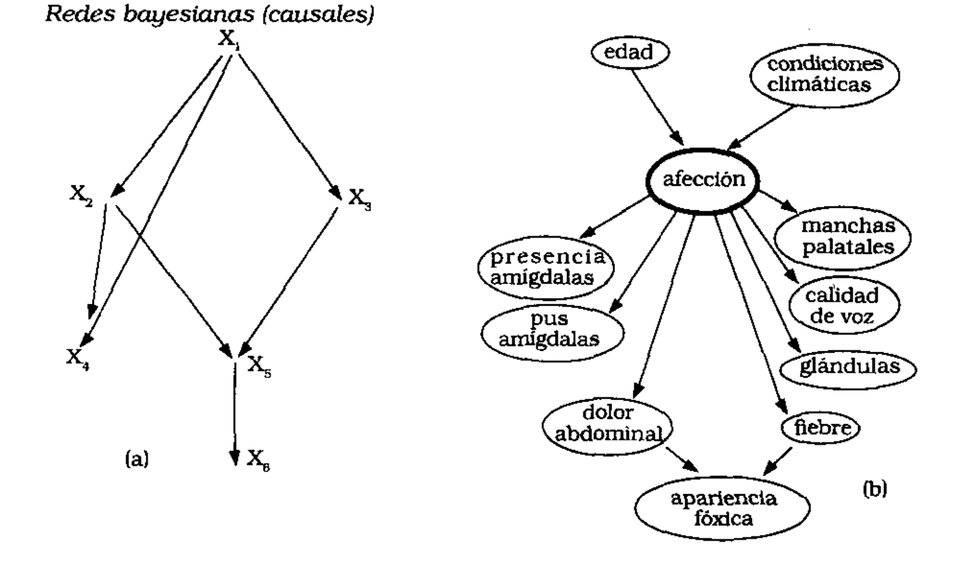
Este valor, junto con los de las ramas que concurren al nodo extremo, determina el valor de k0i, que tendrá una interpretación determinada por su curva de membresía.

Se creó la base RED11 a partir de RED1 con un nodo intermedio v01. Se incluye como conclusión para verificar su aceptabilidad.

Nodos intermedios para REDFUSA



Redes bayesianas (causales)



Finalmente, conviene mencionar las *redes bayesianas*, desarrolladas por Pearl (1987, 1992) a partir, según él mismo menciona, del *análisis de trayectorias* del genetista S. Wright en 1921 y de otros autores de diversos campos: economía (Wold 1964) y sociología (Blalock 1971). Este último autor, Blalock (1985), y otros del mismo grupo crearon los *modelos causales* y han desarrollado distintos modelos en las ciencias sociales. Desde 1969 Blalock ha empleado representaciones en redes para la modelación en ciencias sociales (Blalock, 1969).

En la figura se muestran dos redes, grafos acíclicos dirigidos (GAD). La red (a) es un esquema de Pearl y la (b) uno de la enciclopedia mencionada.

En ambas opera la distribución de probabilidad:

$$P(x_1, x_2, \ldots, x_n) = P(x_n : x_{n-1}, \ldots, x_1) P(x_{n-1} : x_{n-2}, \ldots, x_1) \ldots$$
$$P(x_3 : x_2, x_1) P(X_2 : X_1)$$
$$= \Pi_i P(x_i : \Pi x_i),$$

que es el producto de los valores de probabilidad condicional, para todas las variables:

$$x_1, \ldots, x_n$$

de donde la probabilidad de cada nodo está condicionada por la conjunción de los valores de los nodos padres. Esta conjunción se indica por comas.

Para el ejemplo de Pearl, esquema (a), se tendría:

$$P(x_1, x_2, x_3, x_4, x_5, x_6) =$$
$$P(x_6 : x_5) P(x_5 : x_2, x_3) P(x_4 : x_1 : x_2) P(x_3 : x_1) P(x_2 : x_1) P(x_1)$$

Estas redes implican una considerable simplificación respecto de la aplicación directa del teorema de Bayes. Aun así, se requiere de una cantidad extensa de información previa a su aplicación, Sin embargo, en las redes fusas la situación no es mucho mejor, pues allí se requiere, además, conocer las curvas de membresía en cada nodo.

Como se indica, las redes son causales. No es el caso de los modelos con reglas de producción y grafos, en donde los nodos padre son condiciones, no necesariamente causas. Se puede decir que estas redes son aplicables a los problemas de diagnóstico. Antes se presentaron las aplicaciones PLANDIAG, MEDIAG..., adaptados del trabajo de Naylor.[10] El proceso es muy similar. Aquí se hace una extensión al considerar la representación en red y la conjunción de dos o más condiciones para el cálculo de la probabilidad de una variable

$P(x_i:x_{i-1}, \ldots, x_1)$

Como antes, se prescribe que las condiciones sean independientes, cuestión ya mencionada en el artículo de *Incertidumbre*.

En la *enciclopedia* se menciona la necesidad de hacer una partición para los conceptos de cada nodo. Por ejemplo, para la edad en infante, niño y adulto. Para la fiebre en normal, media y alta. Es, según se advierte, la *discretización* de las variables. No obstante la polémica arriba mencionada, esta discretización se parece mucho al tratamiento lingüístico de los conjuntos fusos, faltaría sólo fijar niveles de probabilidad para cada parámetro.

Estas interpretaciones dan lugar a un elevado número de combinaciones de condiciones para determinar la probabilidad de un evento. Parece ser que para contender con un problema inferencial de diagnóstico se requiere:

- Un acervo extenso de casos bien definidos.
- La interacción externa con un usuario que establezca los niveles de probabilidad: muy alto, bajo, más o menos ... . La partición la proporcionará el propio usuario, mediante *estimaciones* subjetivas.

Hay muchos otros acercamientos al problema inferencial (Neapolitan, 1990). El enfoque de Neapolitan es similar a los anteriores. Básicamente emplea grafos acíclicos dirigidos y la probabilidad bayesiana.

Una última referencia importante es la de Duda, Hart y Nilsson (1976). Se orienta a sistemas de inferencia basados en reglas (de producción).

Nota final

Un aspecto común a los sistemas para abordar el razonamiento aproximado es la extensa cantidad de teoremas y razonamientos matemáticos que pretenden sustentar tal razonamiento. Hay una escasez sustancial de aplicaciones, que puede deberse a las *complicaciones* inherentes o a intereses comerciales.

Notas

1 Un libro con las primeras aportaciones fue recopilado por Findler (1979). En otro texto se recopilan y comparan estos métodos en el contexto de la lógica simbólica (Thayse, 1988).

2 Como en los sistemas de Ch. Naylor y *Expincer* de E. J. Salazar, mencionados en el artículo "Incertidumbre" (Salazar, 1995) Es nuestro propósito la estructuración del conocimiento. En ocasiones estos esquemas se emplean en el desarrollo de *expertos artificiales* orientados sólo a fines comerciales.

3 Mencionadas en Salazar, 1995.

4 Definidas por Shortliffe en *Mycin* (1976), en donde los valores de la certidumbre varían en [-1,+1].

5 Ésta difiere de la expresión de probabilidad

$$p(\text{not } e) = 1 - p(e)$$

6 Por ejemplo Terano, Asai y Sugeno, 1994 y Klir y Folger, 1988.

7 *Una expresión simple para la curva es*

$$S(x) = 1/(1 - e^{-cx})$$

Para otras expresiones ver Cox (1994).

8 Se caracterizan por su simetría respecto a algún valor, como las curvas de distribución normal de probabilidad (curva de error) o la campana de Gauss:

$$G(x;k,\gamma) = \exp(-k(\gamma - x))^2,$$

o para variaciones pronunciadas en el punto más alto:

$$B(x;\gamma,\beta) = 1/(1 + (x - \gamma)/\beta)^2)$$

denominada curva beta. Cox examina un extenso número de curvas y composiciones.

9 Para las relaciones usuales ver el programa *Anis5* (Salazar, 1993a).

10 Sobre este trabajo el autor realizó dos ensayos:

- *Expincer*, que calcula la probabilidad bayesiana concatenada (Salazar, s.f.); programa mencionado en el artículo *Incertidumbre*.
- *Bayes*, que calcula dicha probabilidad, pero anotando los casos desechados que están abajo de un nivel de probabilidad fijado por el usuario (Salazar, 1995a).

Bibliografía

Blalock, H. M.
1969 *Theory construction [From verbal to mathematical formulations]*, Prentice-Hall, Englewood Cliffs, N. J.

Blalock, H. M. Jr.
1985 *Causal models in panel and experimental designs*, Aldine, Nueva York.

Boden, M. A.
1990 *Filosofía de la inteligencia artificial*, Fondo de Cultura Económica, México.

Buchanan, B. G. y E. H. Shortliffe
1984 *Rule-based expert systems (Theory and practice)*, Prentice-Hall, Englewood Cliffs, N.J.

Cox, E.
1994 *The fuzzy systems handbook (A practioner's guide to building, using and maintaining fuzzy systems J.*, Academia Press, Professional, Boston.

Crosson, F. J. y K. M. Sayre
1967 Philosophy and cybernetics (Essays delivered to the Institute for Artificial Intelligence at the University of Notre Dame), Simon and Schuster, Nueva York.

Charniak, E. y D. McDermott
1985 *Introduction to artificial intelligenc*, Read, Addison-Wesley, Massachusetts.

Duda, R. O., P. Hart y N. J. Nilsson
1976 *Subjective bayesian methods for rule-based inference systems*, Menlo Park, Stanford Research Institute (Technical Report 124), California.

Ferrero, J. J.
1975 *La comunicación y los mass media*, Ediciones Mensajero, Bilbao [1971].

Findler, N. V.
1979 *Associative networks. Representation and use of Knowledge by computers*, Academic Press, Nueva York.

Fischler, M. A. y O. Firschein
1987 *Intelligence. The eye, the brain and the computer*, Reading, Addison-Wesley, Massachusetts.

George, F. A.
1979 *Philosophical foundations of cybernetics*, Abacus Press, Kent.

González, A. J. y D. D. Dankel
1993 *The engineering of Knowledge-based systems (Theory and Practice)*, Prentice-Hall, Englewood Cliffs, N.J.

Klir G. J. y T. A. Folger
1988 *Fuzzy sets. Uncertainty and information*, Prentice-Hall, Englewood Cliffs, N.J.

Kosko, B.
1992 *Neural networks and fuzzy systems*, Englewood Cliffs, Prentice-Hall, N.J.
1993 *Fuzzy thinking*, Hyperión, Nueva York.

Kuipers, B.
1994 *Qualitative reasoning (Modeling and simulation with incomplete knowledge)*, MIT Press, Cambridge, Massachusetts.

Marcellus, D. H.
1989 *Expert systems programming in turbo prolog*, Prentice-Hall, Englewood Cliffs, N.J.

Menéndez Pidal, Ramón
1965 *Diccionario Durván de la lengua española*, Durván, Bilbao.

Moles, A.A.
1986 *La creación científica*, Taurus, Madrid.

Nagle, H. T., B. D. Carrol y J. D. Irwin
1975 *An introduction to computer logic*, Prentice-Hall, Englewood Cliffs, N. J.

Neapolitan, R. E.
1990 *Probabilistic reasoning in expert systems (Theory and algorithms)*, Wiley, Nueva York.

Pearl, J.
1987 *Probabilistic reasoning in intelligent system: networks of plausible inference*, Morgan Kaufman, San Mateo, CA.
1992 *Bayesian decision methods*, Wiley Interscience (*Encyclopedia of Artificial Intelligence*, editada por S. C. Stuart), Nueva York.

Salazar R., E. Javier
1990 *Lógica y expertos*, Universidad Autónoma Metropolitana-Iztapalapa (UAM-I) (Colección CSH), México.
1993 "Conocimiento estructurado", en *BOOLE*, Departamento de Economía-UAM-I, México.
1993a *Programa Anis5*, Departamento de Economía-UAM-I, México.
1994 *Programa Grafuz*, Departamento de Economía-UAM-I, México.
1994a *Programa Redecis*, Departamento de Economía-UAM-I, México.
1995 *Incertidumbre*, Departamento de Economía-UAM-I (Serie de investigación, 14), México.
1995a Programa *Bayes*, Departamento de Economía-UAM-I, México.
1995b *Programa Planester*, Departamento de Economía-UAM-I, México.
1995c *Programa Redfusa*, Departamento de Economía-UAM-I, México.
s.f. *Programa Expincer.*

Salazar R., E. Javier, L. Peñalva y R. Leal (comps.)
1993 *Modelación estructurada del conocimiento en las ciencias sociales*, UAM-I/UAM-X, México.

Shortliffe, E. H.
1976 *Computer-based medical consultation: Mycin*, Elsevier, Nueva York.

Sowa, J. F.
1984 *Conceptual structures: information processing in mind and machine*, Reading, Addison-Wesley, Massachusetts.
1991 *Principles of semantic networks (explorations in the representations of knowledge)*, Morgan & Kaufmann, San Mateo, CA.

Terano T., K. Asai y M. Sugeno
1992 *Fuzzy systems Theory and its applications*, Academic Press, Boston.
1994 *Applied fuzzy systems*, Academic Press, Boston.

Thagard, P.
1992 *Conceptual revolutions*, Princeton, Nueva Jersey.

Thayse, A.
1988 *From standard logic to logic programmming (Introducing a logic based approach to artificial intelligence)*, J.Wiley, Nueva York.

Thorton, Ch. y B. Du Boulay
1992 *Artificial intelligence Through search*, Intellect, Oxford.
Ursic Computing
1993 *Fuzzy reasoning with logic and probability*, Wi., 5210 Trafalgar Place, Madison.

APÉNDICE A

Operaciones y definiciones en la lógica de clases

Pertenencia:
$x \in X$ si x es un elemento de X.

Subconjunto
propio
$S \subset X$, si y sólo si todos los elementos de S son elementos de X y $S \neq X$,
impropio
$S \subseteq X$ si los elementos de S son elementos de X y puede ocurrir que $S = X$

Cardinalidad de un conjunto es el número de elementos del conjunto,
si X = {a, b, c, e, f, m} entonces la cardinalidad #X es 6.

Potencia de A o $\wp(A)$ es el conjunto de todos los conjuntos formados con los elementos de A. Por ejemplo
Si A = { 1, 2, 3 } entonces
$\wp(A) = \{\phi, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, A\}$

Notas:

1. El conjunto vacío ϕ, un conjunto sin elementos, es un *subconjunto* de cualquier conjunto,
2. un conjunto es subconjunto de sí mismo,
3. como puede observarse, se definen dos constantes: el universo (1) y el conjunto vacío (ϕ).

Tabla de operadores booleanos
El *complemento* de un conjunto A comprende a los elementos que no están en *A* pero están en el universo del discurso, se denota con A^c o con una barra arriba de la letra: $\bar{A}$.

Involución, complemento del complemento: $(A^c)^c$ o bien $\bar{\bar{A}}$ (doble raya sobre la A),

- conmutatividad: $A \cup B = B \cup A$,
  $A \cap B = B \cap A$
- asociatividad: $(A \cup B) \cup C = A \cup (B \cup C)$,
  $(A \cap B) \cap c = A \cap (B \cap C)$
- distributividad: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- idempotencia: $A \cup A = A$,
  $A \cap A = A$
- absorciones: $A \cup (A \cap B) = A$,
  $A \cap A \cup B) = A$,
  $A \cup (\bar{A} \cap B) = A \cup B$,
  $A \cap (\bar{A} \cup B) = A \cap B$
  $A \cup 1 = 1$ (1: universo),
  $A \cap \phi = \phi$
- identidad: $A \cup \phi = A$, $A \cap 1 = A$
- contradicción: $A \cap \bar{A} = \phi$
- medio excluido: $A \cap \bar{A} = \phi$
- de *Morgan:* $(A \cap D)^c = \bar{A} \cup \bar{D}$,
  $(A \cup E)^c = \bar{A} \cap \bar{E}$
- diferencia: $B\text{-}A = \{x : x \in B \ \& \ x \notin A\}$
- extensiones de la unión y la intersección:
  $\cup_k = A_1 \cup A_2 \cup \ldots \cup A_n$
  $\cap_k = A_1 \cap A_2 \ldots \cap A_n$
  $(K = 1, 2, \ldots, n)$
- operaciones con las funciones de los conjuntos:
  - $f_{\bar{A}}(x) = 1 - f_A(x)$
  - si $\wedge$ es la conjunción: $f_{A \cap B}(x) = f_A(x) \wedge f_B(x)$
  - si $\vee$ es la disyunción: $f_{A \cup B}(x) = f_A(x) \vee f_B(x)$
- si los conjuntos verifican distintas propiedades de *orden:* reflexividad ($A \subset A$), antisimetría ($A \subset B$ & $B \subset A \rightarrow A = B$) y de transitividad ($A \supset B$ & $B \supset A \rightarrow A \subset C$) se define un *POSET*. Bajo otras condiciones (de *idempotencia, aditividad, absorción ...*) se definen distintos tipos de *LATICIAS (todas* dentro del álgebra de boole).

APÉNDICE B

La base RED1 mencionada en el texto es:

imp("s", "n", "k01", "pos", "a01", "muda", "muda" ,1)

```
imp("a", "n", "k01", "pos", "a02", "pos", "a03",
    1)
imp("s", "n", "k02", "pos", "a03", "muda",
    "muda", 1)
imp("a","n","k02","pos","a01","pos","a04",1)
imp("a", "n", "k03", "pos", "a02", "pos", "a04",
    1 )
imp("a","n","k03","pos","a05","pos","a02",1)
imp("s", "n", "k04", "pos", "a04", "muda",
    "muda", 1)
imp("o", "n", "k04", "pos", "a05", "pos", "a06",
    1)
imp("o","n","k05","pos","a07","pos","a10",1)
imp("s","n","k05","pos","a06","muda","muda",1)
imp("a","n","k05","pos","a08","pos","a09",1)

cond_base ("a01")
cond_base ("a02")
cond_base ("a03")
cond_base ("a04")
cond_base ("a05")
cond_base ("a06")
cond_base ("a07")
cond_base ("a08")
cond_base ("a09")
cond_base ("a10")

conclusión ("k01")
conclusión ("k02")
conclusión ("k03")
conclusión ("k04")
conclusión ("k05")
```

Operaciones con los elementos de un grafo

Inclusión $R \subseteq S$: $\mu_R(x,y) \leq \mu_S(x,y)$
Unión $R \cup S$: $\mu_{R \cup S}(x,y) = \mu_R(x,y) \vee \mu_S(x,y)$
Intersección $R \cap S$: $\mu_{R \cap S}(x,y) = \mu_R(x,y) \wedge \mu_S(x,y)$
Complemento $\bar{R}$: $\mu_{\bar{R}}(x,y) = 1 - \mu_R(x,y)$

siendo, como antes: $\vee$= máx, $\wedge$= mín.

Se dan otras definiciones de operación, como en las relaciones no fusas: la reflexividad, simetría, transitividad y otras, que se extienden a los grafos. A partir de estas relaciones, Terano, Asai y Sugeno (1992), establecen diversos criterios para la inferencia.